

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-030657

(43)Date of publication of application : 31.01.2003

(51)Int.Cl.

G06T 7/00

G06K 9/62

G06K 9/68

G10L 15/06

(21)Application number : 2001-212730

(71)Applicant : GLORY LTD

(22)Date of filing : 12.07.2001

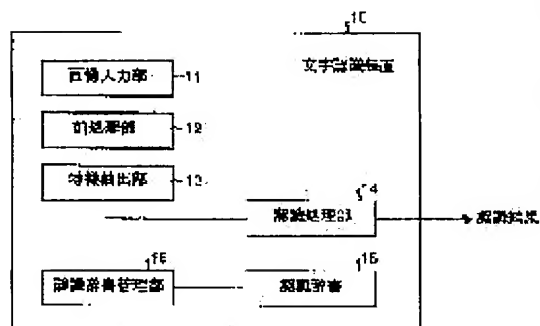
(72)Inventor : KAMEYAMA HIROSHI

(54) PATTERN RECOGNITION DEVICE, PATTERN RECOGNITION METHOD AND PROGRAM MAKING  
COMPUTER EXECUTE THE METHOD

(57)Abstract:

PROBLEM TO BE SOLVED: To improve recognition accuracy in the case of performing non-parametric pattern identification while suppressing a memory capacity for storing a reference pattern even in the case that  $k$  of  $k$  nearest neighbor is  $\geq 3$ .

SOLUTION: The editing processing of a recognition dictionary 15 for eliminating a pattern separated from an identification boundary is performed by a recognition dictionary management part 16 and characters are accurately recognized by high dimensionally projected local linear identification using a Kernel trick by a recognition processing part 14.



\* NOTICES \*

JPO and INPIT are not responsible for any  
damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

## CLAIMS

---

[Claim(s)]

[Claim 1]A pattern recognition device which judges to which category an input pattern characterized by comprising the following belongs based on a recognition dictionary, and performs pattern recognition of said input pattern.

A recognition dictionary which classifies and memorizes two or more reference patterns for every category. An editing means to delete a reference pattern which is distant from a discrimination border between categories among reference patterns in said recognition dictionary.

A recognition means by which partial linearity discernment performs pattern recognition based on a recognition dictionary which deleted a reference pattern which is distant from a discrimination border by said editing means.

[Claim 2]The pattern recognition device according to claim 1, wherein said recognition means performs pattern recognition by partial linearity discernment in high order former space which mapped an original feature vector of an input pattern nonlinear.

[Claim 3]The pattern recognition device according to claim 1 or 2, wherein said recognition means makes a discriminant function a Gaussian kernel which maintains a relation of Euclidean distance in original identification space in high order former space of a map place.

[Claim 4]In a pattern recognition method which judges to which category an input pattern belongs based on a recognition dictionary, and performs pattern recognition of said input pattern, An editing process of deleting a reference pattern which is distant from a discrimination border between categories among reference patterns in a recognition dictionary which classifies and memorizes two or more reference patterns for every category, A pattern recognition method including a recognition step which performs pattern recognition by partial linearity discernment based on a recognition dictionary which deleted a reference pattern which is distant from a discrimination border by said editing process.

[Claim 5]The pattern recognition method according to claim 4, wherein said recognition step performs pattern recognition by partial linearity discernment in high order former space which mapped an original feature vector of an input pattern nonlinear.

[Claim 6]The pattern recognition method according to claim 4 or 5, wherein said recognition step makes a discriminant function a Gaussian kernel which maintains a relation of Euclidean distance in original identification space in high order former space of a map place.

[Claim 7]A program making a computer perform a method indicated to said claims 4-6 and in which computer reading is possible.

## DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Field of the Invention] A pattern recognition device which this invention judges to which category an input pattern belongs based on a recognition dictionary, and performs the pattern recognition of an input pattern, About the program which makes a computer perform a pattern recognition method and a method for the same, especially, even if it is a case where  $k$  by the side of these days  $[k]$  is three or more, It is related with the program which makes a computer perform a pattern recognition device which can raise the recognition precision in the case of performing nonparametric pattern recognition, a pattern recognition method, and a method for the same, controlling the memory space for memorizing a reference pattern.

[0002]

[Description of the Prior Art] Conventionally the thing near an input pattern from a set of sample patterns  $k$  piece selection, The pattern recognition art called  $k$  these days side discernment which determines a classification of an input pattern based on those labels it has is known, and the conventional technology which especially aims at improvement in the speed of processing and improvement in recognition precision these days is known.

[0003] However, since the problem of the storage capacity of a reference pattern and the problem of recognition precision still exist even if it uses such conventional technologies, this applicant, By constituting so that the reference pattern which is distant from the discrimination border between categories among the reference patterns in a recognition dictionary may be deleted in the application for patent No. 347272 [ 2000 to ], It is supposed that the recognition precision in the case of performing nonparametric pattern recognition will be raised, controlling the memory space for memorizing a reference pattern.

[0004] Are a discriminant function based on variable kernel density presumption, and bandwidth  $\sigma_j$  is specifically set as the fixed multiple of these days side distance with a different category, When the dimension of nickel and a pattern is set to  $d$  for the number of reference patterns by the side of these days  $[k]$ , the feature is at the point of omitting  $1/\text{of weighting-factors nickel-}\sigma_j^d$  of the kernel in the strict variable kernel density presuming method.

[0005] Since according to this advanced technology the curved surface to which the middle point of two neighborhood patterns in which categories differ mutually is connected serves as a discrimination border when  $k$  by the side of these days  $[k]$  is 2, the result which may mean that generalization capability becomes high is obtained.

[0006]

[Problem(s) to be Solved by the Invention] However, according to this advanced technology, when  $k$  is three or more, there is a problem whether a desirable result is obtained and that it becomes indefinite. Since it is used by  $k > 2$  in many cases when actually performing character recognition, even if it is a case where it is more than 3 these days side, it is necessary to distinguish a similar character with sufficient accuracy.

[0007] This invention is made in order to solve the problem by the above-mentioned conventional

technology, and it is a thing.

Controlling the memory space for memorizing a reference pattern, even if the purpose is a case where  $k$  is three or more. It is providing the program which makes a computer perform a pattern recognition device which can raise the recognition precision in the case of performing nonparametric pattern recognition, a pattern recognition method, and a method for the same.

[0008]

[Means for Solving the Problem] In order for this invention to solve a technical problem mentioned above and to attain the purpose, a pattern recognition device concerning an invention of claim 1 is characterized by that a pattern recognition device which judges to which category an input pattern belongs based on a recognition dictionary, and performs pattern recognition of said input pattern comprises:

A recognition dictionary which classifies and memorizes two or more reference patterns for every category.

An editing means to delete a reference pattern which is distant from a discrimination border between categories among reference patterns in said recognition dictionary.

A recognition means by which partial linearity discernment performs pattern recognition based on a recognition dictionary which deleted a reference pattern which is distant from a discrimination border by said editing means.

[0009] A pattern recognition device concerning an invention of claim 2 performs pattern recognition in an invention of claim 1 by partial linearity discernment in high order former space where said recognition means mapped an original feature vector of an input pattern nonlinear.

[0010] Let a Gaussian kernel in which a pattern recognition device concerning an invention of claim 3 maintains a relation of Euclidean distance [ in / in said recognition means / original identification space ] in an invention of claim 1 or 2 in high order former space of a map place be a discriminant function.

[0011] A pattern recognition method concerning an invention of claim 4, In a pattern recognition method which judges to which category an input pattern belongs based on a recognition dictionary, and performs pattern recognition of said input pattern, An editing process of deleting a reference pattern which is distant from a discrimination border between categories among reference patterns in a recognition dictionary which classifies and memorizes two or more reference patterns for every category, A recognition step which performs pattern recognition by partial linearity discernment based on a recognition dictionary which deleted a reference pattern which is distant from a discrimination border by said editing process was included.

[0012] A pattern recognition method concerning an invention of claim 5 performs pattern recognition in an invention of claim 4 by partial linearity discernment in high order former space where said recognition step mapped an original feature vector of an input pattern nonlinear.

[0013] A pattern recognition method concerning an invention of claim 6 makes a discriminant function a Gaussian kernel which maintains a relation of Euclidean distance [ in / in said recognition step / original identification space ] in high order former space of a map place in an invention of claim 4 or 5.

[0014] By making a computer perform a method indicated to any one of the claims 4–6, machinery reading of

the program of a program concerning an invention of claim 7 becomes possible, and it can realize any one operation of claims 4-6 by computer this.

[0015]

[Embodiment of the Invention]The suitable embodiment of a program which makes a computer perform a pattern recognition device applied to this invention with reference to an accompanying drawing below, a pattern recognition method, and a method for the same is described in detail. Suppose that the case where this invention is applied to a character reader is shown in this embodiment.

[0016](Composition of a character reader) The composition of the character reader concerning this embodiment is explained first. Drawing 1 is a functional block diagram showing the composition of the character reader concerning this embodiment. The character reader shown in the figure approximates a nonlinear class boundary by adopting the partial linearity discernment mentioned later in a classification hyperplane while performing editing of a dictionary. Partial linearity discernment in the high order former space (infinite dimension) mapped nonlinear is performed by adopting the technique called a kernel trick.

[0017]As shown in the figure, this character reader 10 serves as the image input part 11, the pretreatment part 12, the feature extraction part 13, the recognizing processing part 14, and the recognition dictionary 15 from the recognition dictionary Management Department 16.

[0018]Here, the recognition dictionary of a claim is equivalent to the recognition dictionary 15, the editing means of claim 1 corresponds to the recognition dictionary Management Department 16, and a recognition means corresponds to the recognizing processing part 14.

[0019]The image input parts 11 are input devices, such as a scanner which reads the picture of a character optically, and the image data read by this image input part 11 is outputted to the pretreatment part 12.

[0020]The pretreatment part 12 is a treating part which pretreats the image data received from the image input part 11, and specifically, After carrying out binarization of this with a predetermined threshold, acquiring a binary picture, after carrying out data smoothing of this image data and removing a noise, and starting a character from this binary picture, it is a treating part which normalizes this.

[0021]The feature extraction part 13 is a treating part which extracts characteristic quantity from the normalized alphabetic data in which the pretreatment part 12 pretreated, will carry out the mesh rate of the image data of a character to 5x5 etc., and, specifically, will search for the direction of the outline about each mesh. For example, when the direction of an outline is made into eight directions, a 5x5x8=200 dimension feature space will be formed.

[0022]The recognizing processing part 14 is a treating part which judges to which category an input character belongs, and performs pattern recognition based on the decision result by comparing the characteristic quantity extracted from input characters, such as a handwritten character, with the characteristic quantity in the recognition dictionary 15 prepared beforehand.

[0023]Specifically by this recognizing processing part 14, partial linearity discernment in the high order former space (infinite dimension) mapped nonlinear is performed by adopting a kernel trick. Explanation of this kernel trick and partial linearity discernment is mentioned later.

[0024]The recognition dictionary 15 is a dictionary used for recognition of the input character by the recognizing processing part 14, matches a category for every character and, specifically, memorizes the

characteristic quantity (referred data) of a character for this every category.

[0025]The recognition dictionary Management Department 16 is creation and management of the recognition dictionary 15 a treating part to perform, and specifically, As it makes a bandwidth variable and can set up a discriminant function finely, this recognition dictionary Management Department 16 is raising recognition precision, while it reduces the capacity of the recognition dictionary 14 by performing editing (editing) processing in which the pattern which is distant from a discrimination border is deleted.

[0026]The editing processing by the recognition dictionary Management Department 16 which showed (the concept of editing processing), next drawing 1 is explained concretely. Drawing 2 is an explanatory view showing an example of distribution of the two-dimensional reference pattern belonging to two sorts of categories, and drawing 3 is the explanatory view which established the discrimination border in distribution of the reference pattern shown in drawing 2.

[0027]general -- the discernment technique -- (1) -- the parametric discernment technique and (2) -- being classifiable into the nonparametric discernment technique -- (1) -- to the parametric discernment technique. there are secondary discernment etc. from which the linearity discernment from which a discrimination border serves as a hyperplane, and a discrimination border serve as secondary hypersurface -- (2) -- there are a PAZEN classifier with the discrimination border by which smoothness was carried out to the these days side discernment from which a category is separated by the tattered noy boundary, etc. in a nonparametric discrimination border.

[0028]The reference pattern group which belongs to the category A shown with the small rectangle in a figure as shown in drawing 2, Considering the case where the reference pattern group belonging to the category B shown with a big rectangle in a figure exists, and the reference pattern group of the category A is inserted into the reference pattern group belonging to the category B. The discrimination border or the nonparametric discrimination border by which smooth was carried out of secondary discernment as shown in drawing 3 is formed.

[0029]Thus, although a category is discriminable by using the conventional nonparametric discrimination border, if the conventional discrimination border is using as it is, the number of reference patterns which must be memorized to the recognition dictionary 15 will increase. For this reason, at the recognition dictionary Management Department 16, editing processing is performed and the number of reference patterns is reduced.

[0030]Next, the concept of the editing processing performed by this recognition dictionary Management Department 16 is explained still in detail as compared with a PAZEN classifier. If the frequency function of par ZENU Indaw makes N and a kernel function  $K_d [-]$  and sets [ d dimensional data ] a bandwidth to h for  $s_i$  and a data number, [Equation 1]

$$p(x) = \frac{1}{Nh^d} \sum_{i=1}^N K_d \left[ \frac{(x - s_i)}{h} \right] \quad \dots (1)$$

\*\* -- it becomes like. For this reason, if this  $K_d [-]$  and h are chosen appropriately,  $p_n(x)$  will be converged on the probability density distribution of x.

[0031]The necessary condition in this case becomes  $K_d, \geq 0$  integral  $K_d, [and]$

$dx=1 \lim_{N \rightarrow \infty} h=0 \lim_{N \rightarrow \infty} Nh^d=\infty$ .

[0032] If a regular matrix of  $dx$  is set to  $H$  and an upper type is made more into a general form here, [Equation 2]

$$p(x) = \frac{1}{N|H|} \sum_{i=1}^N K_d[H^{-1}(x - s_i)] \quad \dots (2)$$

It becomes.  $|H|$  shall mean the absolute value of the determinant of  $H$ .

[0033] And when a gauss kernel is used, it is about (1) type, [Equation 3]

$$p(x) = \frac{1}{N} \sum_{i=1}^N \frac{1}{(2\pi h^2)^{d/2}} \exp\left\{-\frac{\|x - s_i\|^2}{2h^2}\right\} \quad \dots (3)$$

It becomes.

[0034] About (2) types [Equation 4]

$$p(x) = \frac{1}{N} \sum_{i=1}^N \frac{1}{(2\pi)^{d/2} h^d |\Sigma_i|^{1/2}} \exp\left\{-\frac{1}{2h^2} (x - s_i)^t \Sigma_i^{-1} (x - s_i)\right\} \quad \dots (4)$$

It becomes. However, sigma is taken as a sample covariance procession.

[0035] And considering the case where a PAZEN classifier is used directly, it is a point estimate of the probability density for every category. [Equation 5]

$$p(x|w_i) = \frac{1}{N_i} \sum_{k=1}^{N_i} \frac{1}{(2\pi)^{d/2} h_i^d |\Sigma_i|^{1/2}} \exp\left\{-\frac{1}{2h_i^2} (x - s_k)^t \Sigma_i^{-1} (x - s_k)\right\} \quad \dots (5)$$

It will be considered as a discriminated result with  $w_i$  used as \*\*\*\*\*.

[0036] Drawing 4 is an explanatory view for explaining a discernment concept at the time of using a PAZEN classifier for one-dimensional data. Data shown by O in a figure The normal distribution N of the an average of 190 standard deviation 30 ( $190, 30^2$ ). Data which is used as data with distribution which mixed the normal distribution N of the an average of 380 standard deviation 30 ( $380, 30^2$ ) by two in eight pairs generated artificially, and is shown by \*\* in a figure, It is considered as artificial data with distribution which mixed the normal distribution N of the an average of 230 standard deviation 60 ( $230, 60^2$ ), and the normal distribution N of the an average of 330 standard deviation 10 ( $330, 10^2$ ) by 6 to 4. The number of data for every category may be ten pieces respectively.

[0037]. And fix standard deviation to a value averaged with the mixing ratio as a preset value of a bandwidth about mixed distribution. That is, when a bandwidth of the category A is set as  $(30 \times 8 + 30 \times 2) / 10 = 30$  and a bandwidth of the category B is set as  $(60 \times 6 + 10 \times 4) / 10 = 40$ , a density function respectively presumed using ten data becomes a curve shown in the figure (a).

[0038] When a bandwidth is fixed for every distribution, namely, a bandwidth of the category A, If a bandwidth of 30 and the category B is set as 10 to data of 60 and (4) to data of (3) to data of 30 and (2) to data of (1), it will become like a curve shown in the figure (b). a case where certain x is given here -- this -- it will be

judged as that to which  $x$  belongs to a big category of a frequency function of  $x$ .

[0039] Thus, although it is discriminable using a PAZEN classifier, according to this PAZEN classifier, a problem that data of big  $N$  is needed graduated by the increase in  $d$  called a curse of a dimension, and a problem that a bandwidth is immobilization arise.

[0040] So, at the recognition dictionary Management Department 16 concerning this embodiment, while erasing  $h_i^d |\sigma_i|^{1/2}$  which is a paragraph of a denominator of (5) types, a measure which makes a bandwidth variable is taken. Although posterior probability is computed based on density presumption by a gauss kernel, specifically, 1/of density normalization paragraph  $h_i^d$  will be then disregarded, using  $\sigma$  common to every category.

[0041] If  $C$  is made into the number of categories, it is the posterior probability of category  $w_i$ , [Equation 6]

$$P(w_i|x) = \frac{p(x|w_i)p(w_i)}{\sum_j^C p(x|w_j)p(w_j)}$$

$$= \frac{\sum_{k \in k \text{ 最近傍}} \frac{1}{h_{ik}^d} \exp\{-\frac{1}{2h_{ik}^2} (x-s_{ik})^t \Sigma^{-1} (x-s_{ik})\}}{\sum_j^C \sum_{k \in k \text{ 最近傍}} \frac{1}{h_{jk}^d} \exp\{-\frac{1}{2h_{jk}^2} (x-s_{jk})^t \Sigma^{-1} (x-s_{jk})\}} \quad \dots (6)$$

It becomes.

[0042] It is about discriminant function  $g_i(x)$  to the reference pattern which carried out editing so that it might leave near a category boundary here. [Equation 7]

$$g_i(x) = \frac{\sum_{k \in k \text{ 最近傍}} \exp\{-\frac{1}{2h_{ik}^2} (x-s_{ik})^t \Sigma^{-1} (x-s_{ik})\}}{\sum_j^C \sum_{k \in k \text{ 最近傍}} \exp\{-\frac{1}{2h_{jk}^2} (x-s_{jk})^t \Sigma^{-1} (x-s_{jk})\}} \quad \dots (7)$$

It carries out.

[0043] Bandwidth  $h_{ik}$  of the  $k$ -th reference pattern of the category  $i$  [Equation 8]

$$h_{ik} = \kappa \min_{i \neq j} |s_{ik} - s_{jm}| \quad \dots (8)$$

It carries out and is set as the fixed multiple of the shortest distance with the whole different category pattern. Under the present circumstances, even if it is  $h_{ik}^d/h_{jk}^d=1$ , the discrimination precision it is higher to disregard  $1/h_{jk}^d$  is acquired.

[0044] Next, editing procedure by the recognition dictionary Management Department 16 which showed drawing 1 is explained. Drawing 5 is a flow chart which shows editing procedure by the recognition dictionary Management Department 16 which showed drawing 1.

[0045] As shown in the figure, at this recognition dictionary Management Department 16, it is considered as set  $B$  [ of a sample chosen ] = {all the samples}, all checked CFLG[ $x$ ]s given to the element  $x$  of  $B$  are made the OFF (OFF), and number of search  $r=10k$  and initialization to set will be performed soon (Step S501).



[0046]Then, if this number  $r$  of neighborhood search is not more than  $k$  about the number  $r$  of search soon as compared with  $k'$  (Step S502) (step S502 denial), processing will be ended as it is, When the number  $r$  of search is more than  $k$  soon, the one sample  $x$  is extracted at random out of what is (Step S502 Affirmation) and  $CFLG[x]=OFF$  of the set  $B$  (Step S503).

[0047]And it is checked whether all near  $[r]$   $x$  are the same as a category of  $x$  (Step S504). An operation which revolution which approximates an envelope of pattern distribution belonging to one category separates from a discrimination border in the direction of an inside of distribution, so that this  $k'$  is large, and smooths a discrimination border will be strengthened.

[0048]As a result, when all near  $[r]$   $x$  are the same as a category of  $x$ , after updating (Step S504 Affirmation) and  $B$  to  $B-\{x\}$ , returning all  $CFLG(s)$  at OFF and setting counted value count to zero, it shifts to (Step S505) and Step S503.

[0049]On the other hand, when at least one near  $[r]$   $x$  is not the same as a category of  $x$ , (Step S504 Denial) and  $CFLG[x]$  are made the one (ON), After \*\*\*\*\*ing counted value count, (Step S506), It investigates whether this counted value count is more than number  $|B|$  of a set (Step S507), and when counted value count is not more than number  $|B|$  of a set, it shifts to (Step S507 Denial) and Step S503.

[0050]On the other hand, when counted value count is more than number  $|B|$  of a set, it is referred to as (Step S507 Affirmation) and  $r=r-\Delta r$ , all  $CFLG(s)$  are returned at OFF, and after setting counted value count to zero, it shifts to (Step S508) and Step S502.

[0051]By performing a series of above-mentioned editing processings, the recognition dictionary Management Department 16 can delete and have a reference pattern which is distant from a discrimination border, and capacity of a recognition dictionary can be reduced.

[0052]Drawing 6 is an explanatory view for explaining a reduction process of a reference pattern by the recognition dictionary Management Department 16. If editing processing is applied using a terminating condition that a pattern with which a category is certainly mutually different is contained, to five  $k'=5$ , i.e., every neighborhood, when 200 samples per each category shown in the figure (a) exist, it will become as it is shown in the figure (b).

[0053]And if editing processing is applied using a terminating condition that a pattern with which a category is certainly mutually different is contained, to four  $k'=4$ , i.e., every neighborhood, it will become as it is shown in the figure (c), If editing processing is similarly applied using a terminating condition of  $k'=3$ , it will become as it is shown in the figure (d).

[0054]As shown in these figures, if this editing processing is performed, a reference pattern near a boundary will remain, but reference patterns of a portion which are distant from a boundary will be reduced.

[0055]Next, an effect of bandwidth change by the recognition dictionary Management Department 16 is explained concretely. Drawing 7 is an explanatory view for explaining an example of an effect of bandwidth change by the recognition dictionary Management Department 16.

[0056]Data shown by O in a figure like drawing 4, The normal distribution  $N$  of the an average of 190 standard deviation 30 ( $190, 30^2$ ). Data which is used as data with distribution which mixed the normal distribution  $N$  of the an average of 380 standard deviation 30 ( $380, 30^2$ ) by two in eight pairs generated artificially, and is shown by \*\* in a figure, It is considered as artificial data with distribution which mixed the

normal distribution  $N$  of the an average of 230 standard deviation 60 ( $230, 60^2$ ), and the normal distribution  $N$  of the an average of 330 standard deviation 10 ( $330, 10^2$ ) by 6 to 4. The number of data for every category may be ten pieces respectively.

[0057]As six errors arise when a bandwidth is fixed about mixed distribution (the category A; 30, the category B; 40), as shown in the figure (a), and shown in the figure (b), It became five errors when a bandwidth was fixed for every distribution (the category A; 30, 30 and the category B60, and 10).

[0058]On the other hand, as shown in the figure (c), when a bandwidth was made into these days side distance with a different category, it became three errors and the error number decreased. As shown in the figure (d), when editing processing was performed, it became the two error number. A reason which the error number reduces in this case is that it can form a discrimination border with two categories which face finely.

[0059]Next, it is two categories and a case where it is simplified k these days two pieces side is explained. A bandwidth shall be a fixed multiple of shortest distance  $\min\|s_i - s_j\|$  (however, categories of  $s_i$  and  $s_j$  differ) with a different category pattern.

[0060]In this case, posterior probability of category  $w_1$  shown by (6) formulas, [Equation 9]

$$P(w_1|x) = \frac{\exp\{-\frac{\|x - s_1\|^2}{2h_1^2}\}}{\exp\{-\frac{\|x - s_1\|^2}{2h_1^2}\} + \frac{h_1^d}{h_2^d} \exp\{-\frac{\|x - s_2\|^2}{2h_2^2}\}} \quad \dots (9)$$

It becomes. It is thought that  $s_1$  and  $s_2$  which are the pairs of a pattern soon are  $h_1 = h_2 = \text{kappa} \|s_i - s_j\|$ .

[0061]This sake, [Equation 10]

$$P(w_1|x) = \frac{\exp\{-\frac{\|x - s_1\|^2}{2h_1^2}\}}{\exp\{-\frac{\|x - s_1\|^2}{2h_1^2}\} + \exp\{-\frac{\|x - s_2\|^2}{2h_2^2}\}} \quad \dots (10)$$

When a next door and  $x$  are the middle point  $(s_1 + s_2)/2$  of  $s_1$  and  $s_2$ , as shown in drawing 8, it is set to  $P(w_1|m) = P(w_2|m) = 1/2$ , and a discrimination border will pass along the middle point of  $s_1$  and  $s_2$ .

[0062](The concept of recognition processing), next the processing concept of the recognizing processing part 14 shown in drawing 1 are explained. This recognizing processing part 14 is performing partial linearity discernment in the high order former space (infinite dimension) mapped nonlinear by adopting the partial linearity discernment which performs linearity discernment using the local reference pattern which carries out the whereabouts near the input data which is a recognition object, and the kernel trick mentioned later.

[0063]Drawing 9 is an explanatory view for explaining partial linearity discernment which this recognizing processing part 14 performs. Here, a reference pattern belonging to the category A is illustrated with a small circle, and a reference pattern belonging to the category B is illustrated with a small rectangular head.

[0064]As shown in the figure, when input data  $x$  used as a recognition object is inputted, a circle of the radius  $r$  centering on this input data  $x$  is considered to be a local domain, and local average  $m_A$  and  $m_B$  for every category are calculated. And a separating hyperplane which divides this local average into two equally

vertically is considered, and it identifies by to which of this separating hyperplane input data  $x$  carries out the whereabouts.

[0065]Specifically, a judgment of this input data  $x$  is faced, [Equation 11]

$$f(x) = (m_A - m_B)^t \left( x - \frac{m_A + m_B}{2} \right) \quad \dots (11)$$

$f(x)$  to say is calculated, and if it is  $f(x) > 0$ , input data  $x$  will judge with the thing belonging to the category A. [0066]Next, a support vector machine and a kernel trick are explained. So that it may be indicated as this support vector machine "for something to be ", Institute of Electronics, Information and Communication Engineers, June, 2000, and pp460-466 with Koji Tsuda and "support vector machine", After mapping a feature vector to high order former space by a certain nonlinear transformation, it is the technique of asking for a hyperplane which separates two categories (class) with linearity quadratic programming. A hyperplane for which it asks makes the maximum quantity of a margin which is the minimum of distance of a hyperplane and a training pattern among those which realize linearity separation, and is excellent in respect of generalization capability. In this support vector machine, the purpose mapped to high order former space is to make linearity separation easy, also when the number of training patterns increases, but. A technique which makes computational complexity small is used by replacing directly inner product calculation of a discriminant function in high order former space after a map with a kernel function, without calculating a map of a feature vector. This is called a kernel trick.

[0067]Drawing 10 is an explanatory view for explaining a concept of a kernel trick. Since the data  $a$  which originally belongs to the category A, and the data  $b$  belonging to the category B cannot be classified in the discrimination border  $L1$  when a discrimination border of the category A and the category B is complicated as shown in the figure, it cannot be judged to which of two categories input data belongs.

[0068]However, if a number of dimension is increased, it will become easy to separate two categories in a straight line. For example, although a category is inseparable in the discrimination border  $L1$  shown in the figure, if the discrimination border  $L2$  made into high order origin is used, the data  $a$  and the data  $b$  are classifiable. Thus, in this kernel trick, it will ask for a hyperplane which separates two categories, after mapping a feature vector to high order former space.

[0069]By the way, since an operation of a kernel is required by the number of a support vector, there is a problem that identification processing takes time in this support vector machine, but. Since it is sufficient if a gauss kernel will be calculated only to a pattern soon even if it uses this kernel trick, since editing is performed as this invention already explained, identification processing can be performed promptly.

[0070]Next, it explains still more concretely about partial linearity discernment using a kernel trick which this recognizing processing part 14 performs. Drawing 11 is an explanatory view for explaining a concept of partial linearity discernment of having used a kernel trick.

[0071]Linearity separation will become impossible when a boundary of a category is dramatically complex, if it is performing partial linearity discernment in an original feature space as shown in drawing 11 (a). For example, although a true category boundary winds like a wave when shown in the figure, since a local discrimination border is a straight line, an appropriate result is not obtained.

[0072] On the other hand, since between the straight lines L3 and L4 will serve as a separator of a category if partial linearity discernment in high order former space which mapped an original feature vector nonlinear using a kernel trick is performed as shown in drawing 11 (b), linearity separation is attained by local linearity discernment.

[0073] Next, a partial linear discriminant function in original feature-space  $R^d$  of  $d$  dimension is explained concretely. However, suppose that covariance-matrix  $\sigma_{m_1}$  of the two categories 1 and 2 and  $\sigma_{m_2}$  are equal, and are a fixed multiple of an identity matrix here.

[0074] Partial linear-discriminant-function  $f_{12}(x)$  in this case, [Equation 12]

$$\begin{aligned} f_{12}(x) &= (m_1 - m_2)^t \left( x - \frac{m_1 + m_2}{2} \right) \\ &= \left( \frac{1}{n_1} \sum_{i=1}^{n_1} x_{1i} - \frac{1}{n_2} \sum_{i=1}^{n_2} x_{2i} \right)^t \left( x - \frac{1}{2n_1} \sum_{i=1}^{n_1} x_{1i} - \frac{1}{2n_2} \sum_{i=1}^{n_2} x_{2i} \right) \\ &\dots (12) \end{aligned}$$

If it is next door and  $f_{12}(x) > 0$ , input data  $x$  will be identified if it belongs to the category 1.

[0075] However, local average  $m_1$  and  $m_2$ . It is the average of the neighborhood pattern which belongs to the categories 1 and 2, respectively, and  $x_{1i}$  ( $i = 1, \dots, n_1$ ) and  $x_{2i}$  ( $i = 1, \dots, n_2$ ) are neighborhood patterns which belong to the categories 1 and 2, respectively, [Equation 13]

$$\|x - x_{1i}\|^2 \leq d_k^2, \|x - x_{2i}\|^2 \leq d_k^2 \quad \dots (13)$$

\*\*\*\*\*.  $d_k$  is  $k$  these days side distance.

[0076] Discriminant function  $f_{12}$  [ in / when the these days side pattern in an original feature space chooses the map which becomes the same also in the high order former space of a map place /  $R^d \phi$  ] ( $\phi(x)$ ), [Equation 14]

$$\begin{aligned} f_{12}(\phi(x)) &= \left( \frac{1}{n_1} \sum_{i=1}^{n_1} \phi(x_{1i}) - \frac{1}{n_2} \sum_{i=1}^{n_2} \phi(x_{2i}) \right)^t \left( \phi(x) - \frac{1}{2n_1} \sum_{i=1}^{n_1} \phi(x_{1i}) - \frac{1}{2n_2} \sum_{i=1}^{n_2} \phi(x_{2i}) \right) \\ &= \frac{1}{n_1} \sum_{i=1}^{n_1} \phi(x)^t \phi(x_{1i}) - \frac{1}{n_2} \sum_{i=1}^{n_2} \phi(x)^t \phi(x_{2i}) - \frac{1}{2n_1^2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_1} \phi(x_{1i})^t \phi(x_{1j}) + \frac{1}{2n_2^2} \sum_{i=1}^{n_2} \sum_{j=1}^{n_2} \phi(x_{2i})^t \phi(x_{2j}) \\ &\dots (14) \end{aligned}$$

It becomes.

[0077] Since this discriminant function is expressed by linear combination of an inner product in  $R^d \phi$ , it can apply the technique of a kernel trick. That is, even if it does not actually perform calculation  $\phi(x)$  of a map to high order origin, only calculation of a real valued function can be managed.

[0078] That is, discriminant function  $f_{12}(\phi(x))$  is by a kernel function which has a relation of  $K(x, y)$   $= \phi(x)^t \phi(y)$ , [Equation 15]

$$\begin{aligned} f_{12}(\phi(x)) &= \left( \frac{1}{n_1} \sum_{i=1}^{n_1} K(x, x_{1i}) - \frac{1}{n_2} \sum_{i=1}^{n_2} K(x, x_{2i}) - \frac{1}{2n_1^2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_1} K(x_{1i}, x_{1j}) + \frac{1}{2n_2^2} \sum_{i=1}^{n_2} \sum_{j=1}^{n_2} K(x_{2i}, x_{2j}) \right) \\ &\dots (15) \end{aligned}$$

It becomes.

[0079]The function called a Gaussian kernel[Equation 16]

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{\sigma^2}\right) \quad \dots (16)$$

It is alike and corresponding phi holds a relation of Euclidean distance in original space also in space of a map place.[0080]Jam,[Equation 17]

$$\begin{aligned} \|x - y_1\|^2 < \|x - y_2\|^2 &\Rightarrow \| \phi(x) - \phi(y_1) \|^2 < \| \phi(x) - \phi(y_2) \|^2 \\ \therefore \| \phi(x) - \phi(y) \|^2 &= \phi(x)^T \phi(x) - 2K(x, y) - \phi(y)^T \phi(y) = 2(1 - \exp(-\frac{\|x - y\|^2}{\sigma^2})) \end{aligned} \quad \dots (17)$$

It \*\*\*\*\*. Therefore, k these days side pattern in an original feature space is k these days side pattern also in the space of a map place. phi corresponding to a Gaussian kernel will be mapped to the space of an infinite dimension.

[0081]The partial linear discriminant function (at the time of  $\sigma_1 \neq \sigma_2$ ) in original feature-space  $R^d$  of d dimension although the detailed explanation is omitted here, [Equation 18]

$$f_{12}(x) = \sum_T^{-1} (m_1 - m_2)^T \left( x - \frac{m_1 + m_2}{2} \right) \quad \dots (18)$$

A method of Fischer (Fisher) shown in \*\*\*\*\* drawing 12 can also be used. However,  $\sigma_T$  is all the covariance matrices of the categories 1 and 2.

[0082]Next, procedure of the recognizing processing part 14 shown in drawing 1 is explained. Drawing 13 is a flow chart which shows procedure of the recognizing processing part 14 shown in drawing 1. If this recognizing processing part 14 inputs feature vector x which is input data as shown in the figure (Step S1301), A these days side [ k pieces ] pattern is looked for from N reference patterns (Step S1302), and it is checked whether all these days side [ k pieces ] patterns belong to the same category  $C_0$  (Step S1303).

[0083]As a result, in belonging to the same category  $C_0$  altogether, it recognizes it as a thing belonging to (Step S1304 Affirmation) and category  $C_0$  (Step S1310). In not belonging to the same category  $C_0$  altogether, on the other hand, (Step S1304 Denial), Top two category  $C_1$  and  $C_2$  are chosen (Step S1305), and a partial discriminant function using an already explained kernel trick is applied (Step S1306).

[0084]When a value of this discriminant function is larger than 0, and (Step S1307 Affirmation), It is recognized as it being category  $C_1$  (Step S1308), and when a value of a discriminant function is not larger than 0, it is recognized as their being (Step S1307 Denial) and category  $C_2$  (Step S1309).

[0085]As mentioned above, while performing editing processing of the recognition dictionary 15 in which a pattern which is distant from a discrimination border by the recognition dictionary Management Department 16 is deleted according to this embodiment, Since it constituted so that the recognizing processing part 14 might perform partial line type discernment using a kernel trick, Recognition precision in a case of

performing nonparametric pattern recognition can be raised controlling memory space for memorizing a reference pattern, even if it is a case where  $k$  by the side of these days  $[k]$  is three or more.

[0086] Although this embodiment showed a case where this invention was applied to a character reader, this invention is not limited to this and can be applied to a various-patterns recognition device which recognizes a pattern. However, the numbers of categories, such as an English character, a number, and katakana, are effective about especially few things. About a Chinese character with many categories, it is effective as a means to realize detailed discrimination processing between specific similar categories.

[0087]

[Effect of the Invention] As explained above, according to the invention of claim 1, the reference pattern which is distant from the discrimination border between categories among the reference patterns in the recognition dictionary which classifies and memorizes two or more reference patterns for every category is deleted, Since it constituted so that partial linearity discernment might perform pattern recognition based on the recognition dictionary which deleted the reference pattern which is distant from a discrimination border, The effect that the pattern recognition device which can raise the recognition precision in the case of performing nonparametric pattern recognition is obtained is done so, controlling the memory space for memorizing a reference pattern, even if it is a case where  $k$  by the side of these days  $[k]$  is three or more.

[0088] Since according to the invention of claim 2 it constituted so that partial linearity discernment in the high order former space which mapped the original feature vector of the input pattern nonlinear might perform pattern recognition, Even if it is a case where the discrimination border is complex, the effect that the pattern recognition device which can be recognized with sufficient accuracy is obtained is done so.

[0089] Since according to the invention of claim 3 it constituted so that the Gaussian kernel which maintains the relation of the Euclidean distance in original identification space in the high order former space of a map place might be made into a discriminant function, The effect that the pattern recognition device which can be recognized efficiently without changing the reference pattern which carries out the whereabouts to the neighborhood before and after a map is obtained is done so.

[0090] According to the invention of claim 4, the reference pattern which is distant from the discrimination border between categories among the reference patterns in the recognition dictionary which classifies and memorizes two or more reference patterns for every category is deleted, Since it constituted so that partial linearity discernment might perform pattern recognition based on the recognition dictionary which deleted the reference pattern which is distant from a discrimination border, The effect that the pattern recognition method which can raise the recognition precision in the case of performing nonparametric pattern recognition is acquired is done so, controlling the memory space for memorizing a reference pattern, even if it is a case where  $k$  by the side of these days  $[k]$  is three or more.

[0091] Since according to the invention of claim 5 it constituted so that partial linearity discernment in the high order former space which mapped the original feature vector of the input pattern nonlinear might perform pattern recognition, Even if it is a case where the discrimination border is complex, the effect that the pattern recognition method which can be recognized with sufficient accuracy is acquired is done so.

[0092] Since according to the invention of claim 6 it constituted so that the Gaussian kernel which maintains the relation of the Euclidean distance in original identification space in the high order former space of a map

place might be made into a discriminant function, The effect that the pattern recognition method which can be recognized efficiently without changing the reference pattern which carries out the whereabouts to the neighborhood before and after a map is acquired is done so.

[0093]According to the invention of claim 7, by making a computer perform the method indicated to any one of the claims 4-6, machinery reading of the program becomes possible and a computer can realize any one operation of claims 4-6 by this.

[Brief Description of the Drawings]

[Drawing 1]It is a functional block diagram showing the composition of the character reader concerning this embodiment of the invention.

[Drawing 2]It is an explanatory view showing an example of distribution of the reference pattern belonging to two sorts of categories.

[Drawing 3]It is the explanatory view which provided an example of the discrimination border in distribution of the reference pattern shown in drawing 2.

[Drawing 4]It is an explanatory view for explaining the discernment concept at the time of using a PAZEN classifier.

[Drawing 5]It is a flow chart which shows the editing procedure by the recognition dictionary Management Department which showed drawing 1.

[Drawing 6]It is an explanatory view for explaining the reduction process of the reference pattern by the recognition dictionary Management Department which showed drawing 1.

[Drawing 7]It is an explanatory view for explaining an example of the effect of the bandwidth change by the recognition dictionary Management Department which showed drawing 1.

[Drawing 8]It is an explanatory view for explaining the case where it is simplified two category and k these days two pieces side.

[Drawing 9]It is an explanatory view for explaining the partial linearity discernment which the recognizing processing part shown in drawing 1 performs.

[Drawing 10]It is an explanatory view for explaining the concept of a kernel trick.

[Drawing 11]It is an explanatory view for explaining the concept of partial linearity discernment of having used the kernel trick.

[Drawing 12]It is an explanatory view for explaining the method of Fischer (Fisher).

[Drawing 13]It is a flow chart which shows the procedure of the recognizing processing part shown in drawing 1.

[Description of Notations]

10 Character reader

11 Image input part

12 Pretreatment part

13 Feature extraction part

14 Recognizing processing part

15 Recognition dictionary

16 Recognition dictionary Management Department  
A and B Category

---

[Translation done.]



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号  
特開2003-30657  
(P2003-30657A)

(43) 公開日 平成15年1月31日 (2003.1.31)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	タームコード* (参考)	
G 0 6 T 7/00	3 0 0	G 0 6 T 7/00	3 0 0 F	5 B 0 6 4
G 0 6 K 9/62	6 4 0	G 0 6 K 9/62	6 4 0 A	5 D 0 1 5
		9/68	E	5 L 0 9 6
G 1 0 L 15/06		G 1 0 L 3/00	5 2 1 J	

審査請求 未請求 請求項の数 7 O L (全 13 頁)

(21) 出願番号 特願2001-212730(P2001-212730)

(22) 出願日 平成13年7月12日 (2001.7.12)

(71) 出願人 000001432

グローリー工業株式会社

兵庫県姫路市下手野1丁目3番1号

(72) 発明者 亀山 博史

兵庫県姫路市下手野一丁目3番1号 グロ

ーリー工業株式会社内

(74) 代理人 100089118

弁理士 酒井 宏明 (外1名)

Fターム(参考) 5B064 DA02 DA23

5D015 GG02 GG03

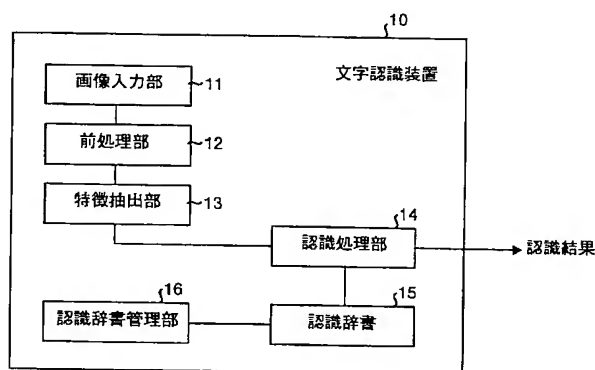
5L096 BA17 JA11 KA03

(54) 【発明の名称】 パターン認識装置、パターン認識方法およびその方法をコンピュータに実行させるプログラム

(57) 【要約】

【課題】 k最近傍のkが3以上の場合であっても、参照パターンを記憶するためのメモリ容量を抑制しつつ、ノンパラメトリックなパターン識別をおこなう場合の認識精度を上げること。

【解決手段】 認識辞書管理部16により識別境界から離れたパターンを削除する認識辞書15のエディティング処理をおこなうとともに、認識処理部14によりカーネルトリックを用いた高次元に射影した局所線型識別により、文字認識を精度良くおこなう。



## 【特許請求の範囲】

【請求項 1】 入力パターンがどのカテゴリーに属するかを認識辞書に基づいて判定して前記入力パターンのパターン認識をおこなうパターン認識装置において、複数の参照パターンをカテゴリーごとに区分して記憶する認識辞書と、

前記認識辞書内の参照パターンのうちカテゴリー間の識別境界から離れた参照パターンを削除するエディティング手段と、

前記エディティング手段により識別境界から離れた参照パターンを削除した認識辞書に基づいて局所線形識別によりパターン認識をおこなう認識手段とを備えたことを特徴とするパターン認識装置。

【請求項 2】 前記認識手段は、入力パターンの原特徴ベクトルを非線形に写像した高次元空間での局所線形識別によりパターン認識をおこなうことを特徴とする請求項 1 に記載のパターン認識装置。

【請求項 3】 前記認識手段は、原識別空間におけるユークリッド距離の関係を写像先の高次元空間で維持するガウシアンカーネルを識別関数とすることを特徴とする請求項 1 または 2 に記載のパターン認識装置。

【請求項 4】 入力パターンがどのカテゴリーに属するかを認識辞書に基づいて判定して前記入力パターンのパターン認識をおこなうパターン認識方法において、複数の参照パターンをカテゴリーごとに区分して記憶する認識辞書内の参照パターンのうちカテゴリー間の識別境界から離れた参照パターンを削除するエディティング工程と、

前記エディティング工程により識別境界から離れた参照パターンを削除した認識辞書に基づいて局所線形識別によりパターン認識をおこなう認識工程とを含んだことを特徴とするパターン認識方法。

【請求項 5】 前記認識工程は、入力パターンの原特徴ベクトルを非線形に写像した高次元空間での局所線形識別によりパターン認識をおこなうことを特徴とする請求項 4 に記載のパターン認識方法。

【請求項 6】 前記認識工程は、原識別空間におけるユークリッド距離の関係を写像先の高次元空間で維持するガウシアンカーネルを識別関数とすることを特徴とする請求項 4 または 5 に記載のパターン認識方法。

【請求項 7】 前記請求項 4～6 に記載された方法をコンピュータに実行させることを特徴とするコンピュータ読み取り可能なプログラム。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】 この発明は、入力パターンがどのカテゴリーに属するかを認識辞書に基づいて判定して入力パターンのパターン認識をおこなうパターン認識装置、パターン認識方法およびその方法をコンピュータに実行させるプログラムに関し、特に、 $k$  最近傍の  $k$  が

3 以上の場合であっても、参照パターンを記憶するためのメモリ容量を抑制しつつ、ノンパラメトリックなパターン識別をおこなう場合の認識精度を上げることができるパターン認識装置、パターン認識方法およびその方法をコンピュータに実行させるプログラムに関する。

## 【0002】

【従来の技術】 従来、サンプルパターンの集合から入力パターンに近いものを  $k$  個選び、それらの持つラベルに基づいて入力パターンの分類を決定する  $k$  最近傍識別と呼ばれるパターン認識技術が知られており、特に最近では、処理の高速化並びに認識精度の向上を図る従来技術が知られている。

【0003】 ところが、これらの従来技術を用いたとしても、参照パターンの記憶容量上の問題や、認識精度上の問題が依然として存在するため、本件出願人は、特願 2000-347272 号において、認識辞書内の参照パターンのうちカテゴリー間の識別境界から離れた参照パターンを削除するよう構成することにより、参照パターンを記憶するためのメモリ容量を抑制しつつ、ノンパラメトリックなパターン識別をおこなう場合の認識精度を上げることとしている。

【0004】 具体的には、可変カーネル密度推定に基づいた識別関数で、バンド幅  $\sigma_j$  を異カテゴリーとの最近傍距離の定数倍に設定しており、 $k$  最近傍の参照パターン数を  $N_i$ 、パターンの次元を  $d$  とした場合に、厳密な可変カーネル密度推定法におけるカーネルの重み係数  $1/N_i \cdot \sigma_j^d$  を省略する点に特徴がある。

【0005】 この先行技術によれば、 $k$  最近傍の  $k$  が 2 のときには、2 個の互いにカテゴリーが異なる近傍パターンの中点を結ぶ曲面が識別境界となるので、汎化能力が高くなるという意味で良い結果が得られる。

## 【0006】

【発明が解決しようとする課題】 しかしながら、この先行技術によれば、 $k$  が 3 以上の場合に望ましい結果が得られるか否かが不明確になるという問題がある。実際に文字認識をおこなう際には、 $k > 2$  で使用されることが多いため、3 最近傍以上の場合であっても精度良く類似文字を判別する必要がある。

【0007】 本発明は、上記従来技術による問題点を解決するためになされたものであり、 $k$  最近傍の  $k$  が 3 以上の場合であっても、参照パターンを記憶するためのメモリ容量を抑制しつつ、ノンパラメトリックなパターン識別をおこなう場合の認識精度を上げることができるパターン認識装置、パターン認識方法およびその方法をコンピュータに実行させるプログラムを提供することを目的とする。

## 【0008】

【課題を解決するための手段】 上述した課題を解決し、目的を達成するため、請求項 1 の発明に係るパターン認識装置は、入力パターンがどのカテゴリーに属するかを

認識辞書に基づいて判定して前記入力パターンのパターン認識をおこなうパターン認識装置において、複数の参照パターンをカテゴリーごとに区分して記憶する認識辞書と、前記認識辞書内の参照パターンのうちカテゴリー間の識別境界から離れた参照パターンを削除するエディティング手段と、前記エディティング手段により識別境界から離れた参照パターンを削除した認識辞書に基づいて局所線形識別によりパターン認識をおこなう認識手段とを備えたことを特徴とする。

【0009】また、請求項2の発明に係るパターン認識装置は、請求項1の発明において、前記認識手段は、入力パターンの原特徴ベクトルを非線形に写像した高次元空間での局所線形識別によりパターン認識をおこなうことを特徴とする。

【0010】また、請求項3の発明に係るパターン認識装置は、請求項1または2の発明において、前記認識手段は、原識別空間におけるユークリッド距離の関係を写像先の高次元空間で維持するガウシアンカーネルを識別関数とすることを特徴とする。

【0011】また、請求項4の発明に係るパターン認識方法は、入力パターンがどのカテゴリーに属するかを認識辞書に基づいて判定して前記入力パターンのパターン認識をおこなうパターン認識方法において、複数の参照パターンをカテゴリーごとに区分して記憶する認識辞書内の参照パターンのうちカテゴリー間の識別境界から離れた参照パターンを削除するエディティング工程と、前記エディティング工程により識別境界から離れた参照パターンを削除した認識辞書に基づいて局所線形識別によりパターン認識をおこなう認識工程とを含んだことを特徴とする。

【0012】また、請求項5の発明に係るパターン認識方法は、請求項4の発明において、前記認識工程は、入力パターンの原特徴ベクトルを非線形に写像した高次元空間での局所線形識別によりパターン認識をおこなうことを特徴とする。

【0013】また、請求項6の発明に係るパターン認識方法は、請求項4または5の発明において、前記認識工程は、原識別空間におけるユークリッド距離の関係を写像先の高次元空間で維持するガウシアンカーネルを識別関数とすることを特徴とする。

【0014】また、請求項7の発明に係るプログラムは、請求項4～6のいずれか一つに記載された方法をコンピュータに実行させることで、そのプログラムを機械読み取り可能となり、これによって、請求項4～6のいずれか一つの動作をコンピュータによって実現することができる。

【0015】

【発明の実施の形態】以下に添付図面を参照して、この発明に係るパターン認識装置、パターン認識方法およびその方法をコンピュータに実行させるプログラムの好適

な実施の形態を詳細に説明する。なお、本実施の形態では、本発明を文字認識装置に適用した場合を示すこととする。

【0016】（文字認識装置の構成）まず、本実施の形態に係る文字認識装置の構成について説明する。図1は、本実施の形態に係る文字認識装置の構成を示す機能ブロック図である。同図に示す文字認識装置は、辞書のエディティングをおこなうとともに、後述する局所線形識別を採用することにより、非線形なクラス境界を区分超平面上で近似することを中心とする。また、カーネルトリックと呼ばれる手法を採用することにより、非線形に写像した高次元空間（無限次元）における局所線形識別をおこなうことを特徴とする。

【0017】同図に示すように、この文字認識装置10は、画像入力部11と、前処理部12と、特徴抽出部13と、認識処理部14と、認識辞書15と、認識辞書管理部16とからなる。

【0018】ここで、請求の範囲の認識辞書は認識辞書15に対応し、請求項1のエディティング手段は認識辞書管理部16に対応し、認識手段は認識処理部14に対応する。

【0019】画像入力部11は、文字の画像を光学的に読み取るスキャナなどの入力デバイスであり、この画像入力部11によって読み取られた画像データは前処理部12に出力される。

【0020】前処理部12は、画像入力部11から受け取った画像データの前処理をおこなう処理部であり、具体的には、この画像データを平滑化処理してノイズを除去した後に、これを所定のしきい値で二値化して二値画像を取得し、この二値画像から文字を切り出した後に、これを正規化する処理部である。

【0021】特徴抽出部13は、前処理部12が前処理をおこなった正規化された文字データから特徴量を抽出する処理部であり、具体的には、文字の画像データを5×5などにメッシュ割りし、各メッシュについての輪郭の方向を求めることになる。たとえば、輪郭の方向を8方向とすると、 $5 \times 5 \times 8 = 200$ 次元の特徴空間が形成されることになる。

【0022】認識処理部14は、手書き文字などの入力文字から抽出された特徴量をあらかじめ用意された認識辞書15内の特徴量と照合することにより、入力文字がどのカテゴリーに属するかを判定し、その判定結果に基づいてパターン認識をおこなう処理部である。

【0023】具体的には、この認識処理部14では、カーネルトリックを採用することにより、非線形に写像した高次元空間（無限次元）における局所線形識別をおこなっている。なお、このカーネルトリックおよび局所線形識別の説明については後述する。

【0024】認識辞書15は、認識処理部14による入力文字の認識に利用する辞書であり、具体的には、各文

字ごとにカテゴリーを対応付け、該カテゴリーごとに文字の特徴量（参照データ）を記憶する。

【0025】認識辞書管理部16は、認識辞書15の作成並びに管理をおこなう処理部であり、具体的には、この認識辞書管理部16は、識別境界から離れたパターンを削除するエディティング（editing）処理をおこなうことにより、認識辞書14の容量を低減するとともに、バンド幅を可変にして識別関数をきめ細かく設定できるようにして認識精度を高めている。

【0026】（エディティング処理の概念）次に、図1に示した認識辞書管理部16によるエディティング処理について具体的に説明する。図2は、2種のカテゴリーに属する2次元の参照パターンの分布の一例を示す説明図であり、図3は、図2に示す参照パターンの分布に識別境界を設けた説明図である。

【0027】一般的に、識別手法は、（1）パラメトリックな識別手法と（2）ノンパラメトリックな識別手法に区分することができ、（1）パラメトリックな識別手法には、識別境界が超平面となる線形識別と識別境界が2次超曲面となる2次識別などがあり、（2）ノンパラ

$$p(x) = \frac{1}{N h^d} \sum_{i=1}^N K_d \left[ \frac{(x - s_i)}{h} \right]$$

のようになる。このため、この $K_d[\cdot]$ と $h$ を適切に選べば、 $p_n(x)$ は $x$ の確率密度分布に収束する。

【0031】この場合の必要条件は、

$$\begin{aligned} K_d[\cdot] &\geq 0 \\ \int K_d[\cdot] dx &= 1 \\ \lim_{N \rightarrow \infty} h &= 0 \end{aligned}$$

$$p(x) = \frac{1}{N |H|} \sum_{i=1}^N K_d[H^{-1}(x - s_i)]$$

となる。なお、 $|H|$ は $H$ の行列式の絶対値を意味するものとする。

【0033】そして、ガウスカーネルを使うと、（1）

$$p(x) = \frac{1}{N} \sum_{i=1}^N \frac{1}{(2\pi h^2)^{d/2}} \exp\left\{-\frac{|x - s_i|^2}{2h^2}\right\}$$

となる。

【0034】また、（2）式については、

$$p(x) = \frac{1}{N} \sum_{i=1}^N \frac{1}{(2\pi)^{d/2} h^d |\Sigma|^{1/2}} \exp\left\{-\frac{1}{2h^2} (x - s_i)^t \Sigma^{-1} (x - s_i)\right\}$$

となる。ただし、 $\Sigma$ は標本共分散行列とする。

【0035】そして、直接パーゼンクラシファイアを使う場合を考えると、カテゴリーごとの確率密度の推定値

もつパーゼンクラシファイアなどがある。

【0028】図2に示すように、図中の小さな矩形で示すカテゴリーAに属する参照パターン群と、図中に大きな矩形で示すカテゴリーBに属する参照パターン群が存在し、カテゴリーAの参照パターン群がカテゴリーBに属する参照パターン群に挟まれている場合を考えると、図3に示すような2次識別の識別境界または平滑されたノンパラメトリックな識別境界が形成される。

【0029】このように、従来のノンパラメトリックな識別境界を用いることにより、カテゴリーの識別をおこなうことができるわけであるが、従来の識別境界をそのまま用いることとすると、認識辞書15に記憶せねばならない参照パターン数が多くなってしまう。このため、認識辞書管理部16では、エディティング処理をおこなって参照パターン数を低減している。

【0030】次に、この認識辞書管理部16によりおこなわれるエディティング処理の概念についてパーゼンクラシファイアと比較してさらに詳細に説明する。パーゼンウインドウの確率密度関数は、 $d$ 次元データを $s_i$ 、データ数を $N$ 、カーネル関数を $K_d[\cdot]$ 、バンド幅を $h$ とすると、

$$\text{【数1】} \quad \dots (1)$$

$$\lim_{N \rightarrow \infty} N h^d = \infty$$

となる。

【0032】ここで、 $d \times d$ の正則行列を $H$ として、上式をより一般形にすると、

【数2】

$$\dots (2)$$

式については、

【数3】

【数4】

である

【数5】

$$p(x|w_i) = \frac{1}{N_i} \sum_{k=1}^{N_i} \frac{1}{(2\pi)^{d/2} h_i^d |\Sigma_i|^{1/2}} \exp\left\{-\frac{1}{2h_i^2} (x - s_k)^t \Sigma_i^{-1} (x - s_k)\right\}$$

… (5)

が最大となる  $w_i$  をもって識別結果とすることになる。

【0036】図4は、1次元データを対象としてパーゼンクラシファイアを用いた場合の識別概念を説明するための説明図である。なお、図中に○で示すデータは、平均190、標準偏差30の正規分布  $N(190, 30^2)$  と、平均380、標準偏差30の正規分布  $N(380, 30^2)$  とを8対2で混合した分布をもつ人工的に発生させたデータとし、図中に□で示すデータは、平均230、標準偏差60の正規分布  $N(230, 60^2)$  と、平均330、標準偏差10の正規分布  $N(330, 10^2)$  とを6対4で混合した分布をもつ人工データとする。また、カテゴリごとのデータの個数は各々10個とする。

【0037】そして、混合分布についてバンド幅の設定値として標準偏差を混合比率で平均した値に固定する、すなわちカテゴリAのバンド幅を、 $(30 \times 8 + 30 \times 2) / 10 = 30$  に設定し、カテゴリBのバンド幅を、 $(60 \times 6 + 10 \times 4) / 10 = 40$  に設定したとき、各々10個のデータを用いて推定される密度関数は、同図(a)に示す曲線になる。

【0038】また、分布ごとにバンド幅を固定すると、すなわちカテゴリAのバンド幅を、(1)のデータに

$$P(w_i|x) = \frac{p(x|w_i)p(w_i)}{\sum_j^C p(x|w_j)p(w_j)}$$

$$\approx \frac{\sum_{k \in k_{\text{最近傍}}} \frac{1}{h_{ik}^d} \exp\left\{-\frac{1}{2h_{ik}^2} (x - s_{ik})^t \Sigma^{-1} (x - s_{ik})\right\}}{\sum_j^C \sum_{k \in k_{\text{最近傍}}} \frac{1}{h_{jk}^d} \exp\left\{-\frac{1}{2h_{jk}^2} (x - s_{jk})^t \Sigma^{-1} (x - s_{jk})\right\}}$$

… (6)

となる。

【0042】また、ここではカテゴリ境界付近のみを残すようにエディティングした参照パターンに対して、

$$g_i(x) = \frac{\sum_{k \in k_{\text{最近傍}}} \exp\left\{-\frac{1}{2h_{ik}^2} (x - s_{ik})^t \Sigma^{-1} (x - s_{ik})\right\}}{\sum_j^C \sum_{k \in k_{\text{最近傍}}} \exp\left\{-\frac{1}{2h_{jk}^2} (x - s_{jk})^t \Sigma^{-1} (x - s_{jk})\right\}}$$

… (7)

とする。

【0043】また、カテゴリ  $i$  の  $k$  番目の参照パターン

$$h_{ik} = \kappa \min_{j \neq i} |s_{ik} - s_{jm}|$$

… (8)

として、異カテゴリパターン全体との最小距離の定数倍に設定する。この際、 $h_{ik}^d / h_{jk}^d \neq 1$  であっても、 $1/h_{jk}^d$  を無視した方が高い識別精度が得られる。

【0044】次に、図1に示した認識辞書管理部16に

対して30および(2)のデータに対して30、カテゴリBのバンド幅を、(3)のデータに対して60および(4)のデータに対して10に設定すると、同図

(b)に示す曲線のようにになる。ここで、ある  $x$  が与えられた場合に、該  $x$  の確率密度関数の大きなカテゴリに  $x$  が属するものと判断されることになる。

【0039】このように、パーゼンクラシファイアを使って識別をおこなうことができるわけであるが、かかるパーゼンクラシファイアによれば、次元の呪いと呼ばれる  $d$  の増加によって累進的に大きな  $N$  のデータが必要になるという問題や、バンド幅が固定であるという問題が生ずる。

【0040】そこで、本実施の形態にかかる認識辞書管理部16では、(5)式の分母の項である  $h_i^d |\Sigma_i|^{1/2}$  を消すとともに、バンド幅を可変にする対策を講じている。具体的には、ガウスカーネルによる密度推定に基づいて事後確率を算出するが、そのときどのカテゴリにも共通の  $\Sigma$  を用い、かつ、密度正規化項  $1/h_i^d$  を無視することにする。

【0041】 $C$  をカテゴリ数とすると、カテゴリ  $w_i$  の事後確率は、

【数6】

識別関数  $g_i(x)$  を

【数7】

ンのバンド幅  $h_{ik}$  を

【数8】

よるエディティング処理手順について説明する。図5は、図1に示した認識辞書管理部16によるエディティング処理手順を示すフローチャートである。

【0045】同図に示すように、この認識辞書管理部1

6では、選択されるサンプルの集合 $B = \{\text{全サンプル}\}$ とし、 $B$ の要素 $x$ に付与されたチェック済み $CFLG[x]$ をすべてオフ(OFF)にし、近傍探索数 $r = 10k$ とセットする初期化をおこなう(ステップS501)。

【0046】その後、近傍探索数 $r$ を $k'$ と比較して(ステップS502)、この近傍探索数 $r$ が $k$ 以上でなければ(ステップS502否定)そのまま処理を終了し、近傍探索数 $r$ が $k$ 以上である場合には(ステップS502肯定)、集合 $B$ の $CFLG[x] = OFF$ であるもののの中から、ランダムにサンプル $x$ を1個抽出する(ステップS503)。

【0047】そして、 $x$ の近傍 $r$ 個の全部が $x$ のカテゴリと同じであるか否かを確認する(ステップS504)。なお、この $k'$ が大きいほど1つのカテゴリに属するパターン分布の外皮を近似する回帰が識別境界から分布の内部の方向に離れ、また識別境界を平滑化する作用を強めることになる。

【0048】その結果、 $x$ の近傍 $r$ 個の全部が $x$ のカテゴリと同じである場合には(ステップS504肯定)、 $B$ を $B - \{x\}$ に更新し、 $CFLG$ をすべてOFFに戻してカウント値countを0とした後に(ステップS505)、ステップS503に移行する。

【0049】これに対して、 $x$ の近傍 $r$ 個の一つでも $x$ のカテゴリと同じでない場合には(ステップS504否定)、 $CFLG[x]$ をオン(ON)にして、カウント値countをインクリメントした後に(ステップS506)、このカウント値countが集合の個数 $|B|$ 以上であるか否かを調べ(ステップS507)、カウント値countが集合の個数 $|B|$ 以上でない場合には(ステップS507否定)、ステップS503に移行する。

【0050】一方、カウント値countが集合の個数 $|B|$ 以上である場合には(ステップS507肯定)、 $r = r - \Delta r$ とし、 $CFLG$ をすべてOFFに戻し、カウント値countを0とした後に(ステップS508)、ステップS502に移行する。

【0051】上記一連のエディティング処理をおこなうことにより、認識辞書管理部16が識別境界から離れた参照パターンを削除し、もって認識辞書の容量を低減することができる。

【0052】図6は、認識辞書管理部16による参照パターンの削減過程を説明するための説明図である。同図(a)に示す各カテゴリあたり200個のサンプルが存在する場合に、 $k' = 5$ すなわち、どの近傍5個についても必ず互いにカテゴリの違うパターンが含まれているという終了条件を用いてエディティング処理を適用

すると、同図(b)のようになる。

【0053】そして、 $k' = 4$ すなわち、どの近傍4個についても必ず互いにカテゴリの違うパターンが含まれているという終了条件を用いてエディティング処理を適用すると同図(c)のようになり、同様に $k' = 3$ という終了条件を用いてエディティング処理を適用すると同図(d)のようになる。

【0054】これらの図から分かるように、このエディティング処理をおこなうと、境界付近の参照パターンが残るが、境界から離れた部分の参照パターンは削減されることになる。

【0055】次に、認識辞書管理部16によるバンド幅変更の効果について具体的に説明する。図7は、認識辞書管理部16によるバンド幅変更の効果の一例を説明するための説明図である。

【0056】なお、図4と同様に、図中に○で示すデータは、平均190、標準偏差30の正規分布 $N(190, 30^2)$ と、平均380、標準偏差30の正規分布 $N(380, 30^2)$ とを8対2で混合した分布をもつ人工的に発生させたデータとし、図中に□で示すデータは、平均230、標準偏差60の正規分布 $N(230, 60^2)$ と、平均330、標準偏差10の正規分布 $N(330, 10^2)$ とを6対4で混合した分布をもつ人工データとする。また、カテゴリごとのデータの個数は各々10個とする。

【0057】同図(a)に示すように、混合分布についてバンド幅を固定(カテゴリA; 30、カテゴリB; 40)した場合にはエラーが6個生じ、同図(b)に示すように、分布ごとにバンド幅を固定(カテゴリA; 30と30、カテゴリB 60と10)すると、エラーが5個となった。

【0058】これに対して、同図(c)に示すように、バンド幅を異カテゴリとの最近傍距離とした場合には、エラーが3個となり、エラー個数が低減した。さらに、同図(d)に示すようにエディティング処理をおこなうと、エラー個数が2個となった。かかる場合にエラー個数が低減する理由は、2個の相対するカテゴリとの識別境界をきめ細かく形成できるからである。

【0059】次に、2カテゴリであり、かつ、 $k$ 最近傍が2個と単純化した場合について説明する。なお、バンド幅は、異カテゴリパターンとの最小距離 $\min ||s_i - s_j||$  (ただし、 $s_i$ と $s_j$ のカテゴリは異なる)の定数倍であるものとする。

【0060】この場合には、(6)式で示したカテゴリ $w_i$ の事後確率は、

【数9】

$$P(w_1|x) = \frac{\exp\{-\frac{\|x-s_1\|^2}{2h_1^2}\}}{\exp\{-\frac{\|x-s_1\|^2}{2h_1^2}\} + \frac{h_1^d}{h_2^d} \exp\{-\frac{\|x-s_2\|^2}{2h_2^2}\}} \quad \dots (9)$$

となる。また、近傍パターンのペアである  $s_1$  と  $s_2$  は、  
 $h_1 = h_2 = \kappa \|s_1 - s_2\|$  となっていると考えられる。

【0061】このため、  
 【数10】

$$P(w_1|x) = \frac{\exp\{-\frac{\|x-s_1\|^2}{2h_1^2}\}}{\exp\{-\frac{\|x-s_1\|^2}{2h_1^2}\} + \exp\{-\frac{\|x-s_2\|^2}{2h_2^2}\}} \quad \dots (10)$$

となり、 $x$  が  $s_1$  と  $s_2$  の中点  $(s_1 + s_2) / 2$  のとき、  
 図8に示すように、 $P(w_1|m) = P(w_2|m) = 1/2$  となり、識別境界が  $s_1$  と  $s_2$  の中点を通ることになる。

【0062】（認識処理の概念）次に、図1に示した認識処理部14の処理概念について説明する。この認識処理部14は、認識対象である入力データの近傍に所在する局所的な参照パターンを用いて線形識別をおこなう局所線形識別と、後述するカーネルトリックとを採用することにより、非線形に写像した高次元空間（無限次元）における局所線形識別をおこなっている。

【0063】図9は、この認識処理部14がおこなう局所線形識別を説明するための説明図である。ここでは、

$$f(x) = (m_A - m_B)^t \left( x - \frac{m_A + m_B}{2} \right) \quad \dots (11)$$

という  $f(x)$  を求め、もし  $f(x) > 0$  であれば、入力データ  $x$  がカテゴリ A に属するものと判定することになる。

【0066】次に、サポートベクターマシンおよびカーネルトリックについて説明する。このサポートベクターマシンとは、「津田宏治，“サポートベクターマシンとは何か”，電子情報通信学会誌，2000年6月，pp460-466」に記載されるように、ある非線形変換によって特徴ベクトルを高次元空間に写像した後に、2つのカテゴリ（クラス）を分離する超平面を線形2次計画法により求める手法である。求める超平面は、線形分離を実現するもののうち、超平面と訓練パターンの距離の最小値であるマージンの量を最大にするものであり、汎化能力の点で優れたものである。このサポートベクターマシンにおいて、高次元空間に写像する目的は、訓練パターン数が増加したときにも線形分離を容易にすることであるが、写像後の高次元空間における識別関数の内積計算を、特徴ベクトルの写像を計算せずに直接カーネル関数によって置き換えることによって、計算量を小さくするテクニックが使われている。これがカーネルトリックと呼ばれている。

カテゴリ A に属する参照パターンを小さな円で図示し、カテゴリ B に属する参照パターンを小さな四角で図示している。

【0064】同図に示すように、認識対象となる入力データ  $x$  が入力された場合には、この入力データ  $x$  を中心とする半径  $r$  の円を局所領域と考え、各カテゴリごとの局所平均  $m_A$  および  $m_B$  を求める。そして、この局所平均を垂直に2等分する分離超平面を考え、入力データ  $x$  がこの分離超平面のどちらに所在するかによって識別をおこなう。

【0065】具体的には、この入力データ  $x$  の判定に際しては、

【数11】

【0067】図10は、カーネルトリックの概念を説明するための説明図である。同図に示すように、カテゴリ A とカテゴリ B の識別境界が複雑である場合には、本来カテゴリ A に属するデータ  $a$  とカテゴリ B に属するデータ  $b$  を識別境界  $L_1$  で区分できないため、入力データが2つのカテゴリのどちらに属するかを判断できない。

【0068】しかしながら、次元数を増やすと、2つのカテゴリを直線で分離しやすくなる。たとえば、同図に示す識別境界  $L_1$  ではカテゴリを分離することができないが、高次元にした識別境界  $L_2$  を用いると、データ  $a$  とデータ  $b$  を区分することができる。このように、かかるカーネルトリックでは、特徴ベクトルを高次元空間に写像した後に、2つのカテゴリを分離する超平面を求めることになる。

【0069】ところで、このサポートベクターマシンには、サポートベクターの個数分だけカーネルの演算が必要のために識別処理に時間を要するという問題があるが、本発明ではすでに説明したようにエディティングをおこなっているので、このカーネルトリックを用いたとしても、近傍パターンに対してのみガウスカーネルの計

算をおこなえば足りるので、迅速に識別処理をおこなうことができる。

【0070】次に、この認識処理部14がおこなうカーネルトリックを利用した局所線形識別についてさらに具体的に説明する。図11は、カーネルトリックを利用した局所線形識別の概念を説明するための説明図である。

【0071】図11(a)に示すように、原特徴空間での局所線形識別をおこなうこととすると、カテゴリの境界が非常に入り組んでいる場合に、線形分離が不可能になってしまう。たとえば、同図に示す場合には、真の

10 カテゴリ境界が波のようになっているにもかかわらず、局所的な識別境界が直線であるため、妥当な結果が得られない。

$$f_{12}(x) = (m_1 - m_2)^t \left( x - \frac{m_1 + m_2}{2} \right) \\ = \left( \frac{1}{n_1} \sum_{i=1}^{n_1} x_{1i} - \frac{1}{n_2} \sum_{i=1}^{n_2} x_{2i} \right)^t \left( x - \frac{1}{2n_1} \sum_{i=1}^{n_1} x_{1i} - \frac{1}{2n_2} \sum_{i=1}^{n_2} x_{2i} \right)$$

... (12)

となり、 $f_{12}(x) > 0$ であれば、入力データ $x$ をカ

20 ゴリー1に属すると識別する。

【0075】ただし、局所平均 $m_1$ 、 $m_2$ は、それぞれカテゴリ1および2に属する近傍パターンの平均であ

$$\|x - x_{1i}\|^2 \leq d_k^2, \|x - x_{2i}\|^2 \leq d_k^2$$

... (13)

を満たす。なお、 $d_k$ は $k$ 最近傍距離である。

【0076】また、原特徴空間での最近傍パターンが、写像先の高次元空間においても同一となる写像を選ぶこ

$$f_{12}(\phi(x)) = \left( \frac{1}{n_1} \sum_{i=1}^{n_1} \phi(x_{1i}) - \frac{1}{n_2} \sum_{i=1}^{n_2} \phi(x_{2i}) \right)^t \left( \phi(x) - \frac{1}{2n_1} \sum_{i=1}^{n_1} \phi(x_{1i}) - \frac{1}{2n_2} \sum_{i=1}^{n_2} \phi(x_{2i}) \right) \\ = \frac{1}{n_1} \sum_{i=1}^{n_1} \phi(x)^t \phi(x_{1i}) - \frac{1}{n_2} \sum_{i=1}^{n_2} \phi(x)^t \phi(x_{2i}) - \frac{1}{2n_1^2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_1} \phi(x_{1i})^t \phi(x_{1j}) + \frac{1}{2n_2^2} \sum_{i=1}^{n_2} \sum_{j=1}^{n_2} \phi(x_{2i})^t \phi(x_{2j})$$

... (14)

となる。

【0077】この識別関数は、 $R^d \phi$ において内積の線形和によって表されているので、カーネルトリックの手法を適用できる。つまり、高次元への写像の計算 $\phi$

( $x$ )を実際におこなわなくても実数値関数の計算だけ

$$f_{12}(\phi(x)) = \left( \frac{1}{n_1} \sum_{i=1}^{n_1} K(x, x_{1i}) - \frac{1}{n_2} \sum_{i=1}^{n_2} K(x, x_{2i}) \right) - \frac{1}{2n_1^2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_1} K(x_{1i}, x_{1j}) + \frac{1}{2n_2^2} \sum_{i=1}^{n_2} \sum_{j=1}^{n_2} K(x_{2i}, x_{2j})$$

... (15)

となる。

【0079】ガウシアンカーネルと呼ばれる関数

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{\sigma^2}\right)$$

... (16)

に対応する $\phi$ は、原空間におけるユークリッド距離の関

50 係を、写像先の空間においても保持する

【0072】これに対して、図11(b)に示すように、カーネルトリックを利用して原特徴ベクトルを非線形に写像した高次元空間での局所線形識別をおこなうと、直線 $l_3$ と $l_4$ の間がカテゴリの分離帯となるため、局所的な線形識別によって線形分離が可能となる。

【0073】次に、 $d$ 次元の原特徴空間 $R^d$ における局所線形識別関数について具体的に説明する。ただし、ここでは2カテゴリ1、2の共分散行列 $\Sigma_1$ 、 $\Sigma_2$ が等しく、かつ、単位行列の定数倍であるとする。

【0074】この場合の局所線形識別関数 $f_{12}(x)$ は、

【数12】

とにより、 $R^d \phi$ における識別関数 $f_{12}(\phi(x))$ は、

【数14】

で済ませることができる。

【0078】すなわち、 $K(x, y) = \phi(x)^t \phi(y)$ の関係を持つカーネル関数によって、識別関数 $f_{12}(\phi(x))$ は、

【数15】



【0080】つまり、

$$\|x - y_1\|^2 < \|x - y_2\|^2 \text{ ならば、} \|\phi(x) - \phi(y_1)\|^2 < \|\phi(x) - \phi(y_2)\|^2 \quad \text{【数17】}$$

$$\because \|\phi(x) - \phi(y)\|^2 = \phi(x)^t \phi(x) - 2K(x, y) - \phi(y)^t \phi(y) = 2(1 - \exp(-\frac{\|x - y\|^2}{\sigma^2}))$$

… (17)

が成立する。したがって原特徴空間におけるk最近傍パターンは、写像先の空間においてもk最近傍パターンとなっている。なお、ガウシアンカーネルに対応するφは、無限次元の空間に写像することになる。

$$f_{12}(x) = \sum_T^{-1} (m_1 - m_2)^t (x - \frac{m_1 + m_2}{2})$$

を用いた図12に示すフィッシャー (Fisher) の方法を用いることもできる。ただし、Σ<sub>T</sub>は、カテゴリ1および2の全共分散行列である。

【0082】次に、図1に示した認識処理部14の処理手順について説明する。図13は、図1に示した認識処理部14の処理手順を示すフローチャートである。同図に示すように、この認識処理部14は、入力データである特徴ベクトルxを入力すると (ステップS1301)、N個の参照パターンからk個の最近傍パターンを探し (ステップS1302)、k個の最近傍パターンが全て同一のカテゴリC<sub>0</sub>に属するか否かを確認する (ステップS1303)。

【0083】その結果、全て同じカテゴリC<sub>0</sub>に属する場合には (ステップS1304肯定)、カテゴリC<sub>0</sub>に属するものと認識する (ステップS1310)。これに対して、全て同じカテゴリC<sub>0</sub>に属するわけではない場合には (ステップS1304否定)、上位2つのカテゴリC<sub>1</sub>とC<sub>2</sub>を選択し (ステップS1305)、すでに説明したカーネルトリックを用いた局所識別関数を適用する (ステップS1306)。

【0084】そして、この識別関数の値が0よりも大きい場合には (ステップS1307肯定)、カテゴリC<sub>1</sub>であると認識し (ステップS1308)、識別関数の値が0よりも大きくない場合には (ステップS1307否定)、カテゴリC<sub>2</sub>であると認識する (ステップS1309)。

【0085】上述してきたように、本実施の形態によれば、認識辞書管理部16により識別境界から離れたパターンを削除する認識辞書15のエディティング処理をおこなうとともに、認識処理部14によりカーネルトリックを用いた局所線型識別をおこなうよう構成したので、k最近傍のkが3以上の場合であっても、参照パターンを記憶するためのメモリ容量を抑制しつつ、ノンパラメトリックなパターン識別をおこなう場合の認識精度を上げることができる。

【0086】なお、本実施の形態では、本発明を文字認識装置に適用した場合を示したが、本発明はこれに限定されるものではなく、パターンの認識をおこなう各種パ

【0081】また、ここではその詳細な説明を省略するが、d次元の原特徴空間R<sup>d</sup>における局所線形識別関数 (Σ<sub>1</sub> ≠ Σ<sub>2</sub>のとき) は、

【数18】

… (18)

ターン認識装置に適用することができる。ただし、英文字、数字、カタカナなどカテゴリ数が少ないものについて特に有効である。また、カテゴリ数の多い漢字については、特定の類似するカテゴリ間の詳細判別処理を実現する手段として有効である。

【0087】

【発明の効果】以上説明したように、請求項1の発明によれば、複数の参照パターンをカテゴリごとに区分して記憶する認識辞書内の参照パターンのうちカテゴリ間の識別境界から離れた参照パターンを削除し、識別境界から離れた参照パターンを削除した認識辞書に基づいて局所線形識別によりパターン認識をおこなうよう構成したので、k最近傍のkが3以上の場合であっても、参照パターンを記憶するためのメモリ容量を抑制しつつ、ノンパラメトリックなパターン識別をおこなう場合の認識精度を上げることが可能なパターン認識装置が得られるという効果を奏する。

【0088】また、請求項2の発明によれば、入力パターンの原特徴ベクトルを非線形に写像した高次元空間での局所線形識別によりパターン認識をおこなうよう構成したので、識別境界が入り組んでいる場合であっても精度良く認識することが可能なパターン認識装置が得られるという効果を奏する。

【0089】また、請求項3の発明によれば、原識別空間におけるユークリッド距離の関係を写像先の高次元空間で維持するガウシアンカーネルを識別関数とするよう構成したので、写像の前後で近傍に所在する参照パターンを変えずに効率良く認識することが可能なパターン認識装置が得られるという効果を奏する。

【0090】また、請求項4の発明によれば、複数の参照パターンをカテゴリごとに区分して記憶する認識辞書内の参照パターンのうちカテゴリ間の識別境界から離れた参照パターンを削除し、識別境界から離れた参照パターンを削除した認識辞書に基づいて局所線形識別によりパターン認識をおこなうよう構成したので、k最近傍のkが3以上の場合であっても、参照パターンを記憶するためのメモリ容量を抑制しつつ、ノンパラメトリックなパターン識別をおこなう場合の認識精度を上げるこ

とが可能なパターン認識方法が得られるという効果を奏する。

【0091】また、請求項5の発明によれば、入力パターンの原特徴ベクトルを非線形に写像した高次元空間での局所線形識別によりパターン認識をおこなうよう構成したので、識別境界が入り組んでいる場合であっても精度良く認識することが可能なパターン認識方法が得られるという効果を奏する。

【0092】また、請求項6の発明によれば、原識別空間におけるユークリッド距離の関係を写像先の高次元空間で維持するガウシアンカーネルを識別関数とするよう構成したので、写像の前後で近傍に所在する参照パターンを変えずに効率良く認識することが可能なパターン認識方法が得られるという効果を奏する。

【0093】また、請求項7の発明によれば、請求項4～6のいずれか一つに記載された方法をコンピュータに実行させることで、そのプログラムを機械読み取り可能となり、これによって、請求項4～6のいずれか一つの動作をコンピュータによって実現することができる。

【図面の簡単な説明】

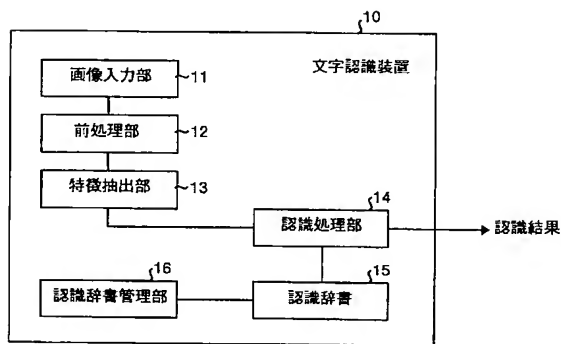
【図1】この発明の実施の形態に係る文字認識装置の構成を示す機能ブロック図である。

【図2】2種のカテゴリに属する参照パターンの分布の一例を示す説明図である。

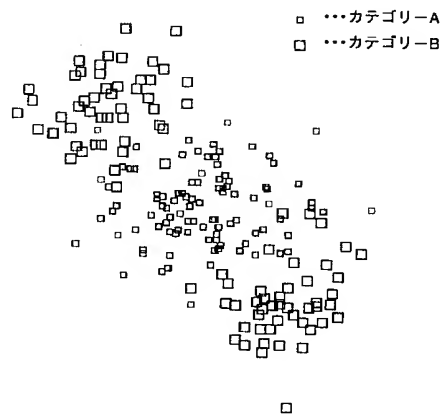
【図3】図2に示す参照パターンの分布に識別境界の一例を設けた説明図である。

【図4】パーゼンクラシファイアを用いた場合の識別概念を説明するための説明図である。

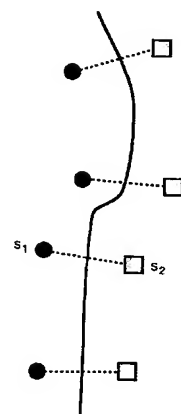
【図1】



【図2】



【図8】



【図5】図1に示した認識辞書管理部によるエディティング処理手順を示すフローチャートである。

【図6】図1に示した認識辞書管理部による参照パターンの削減過程を説明するための説明図である。

【図7】図1に示した認識辞書管理部によるバンド幅変更の効果の一例を説明するための説明図である。

【図8】2カテゴリ、かつ、k最近傍が2個と単純化した場合を説明するための説明図である。

【図9】図1に示した認識処理部がおこなう局所線形識別を説明するための説明図である。

【図10】カーネルトリックの概念を説明するための説明図である。

【図11】カーネルトリックを利用した局所線形識別の概念を説明するための説明図である。

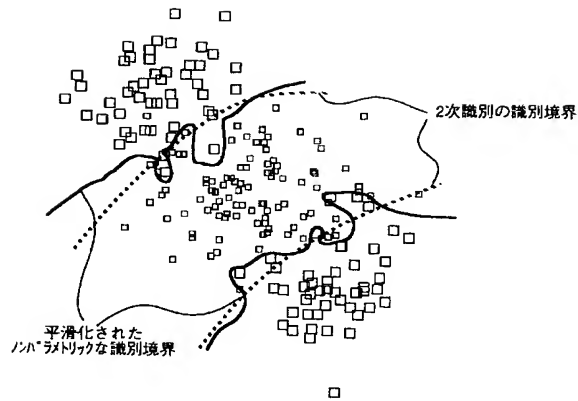
【図12】フィッシャー（Fisher）の方法を説明するための説明図である。

【図13】図1に示した認識処理部の処理手順を示すフローチャートである。

【符号の説明】

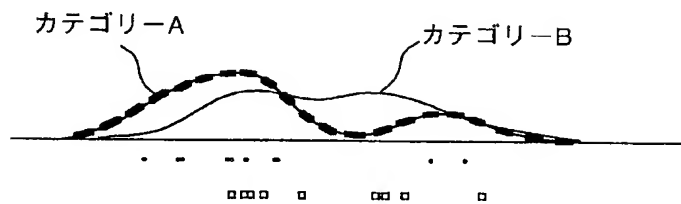
- 10 文字認識装置
- 11 画像入力部
- 12 前処理部
- 13 特徴抽出部
- 14 認識処理部
- 15 認識辞書
- 16 認識辞書管理部
- A, B カテゴリ

【図3】

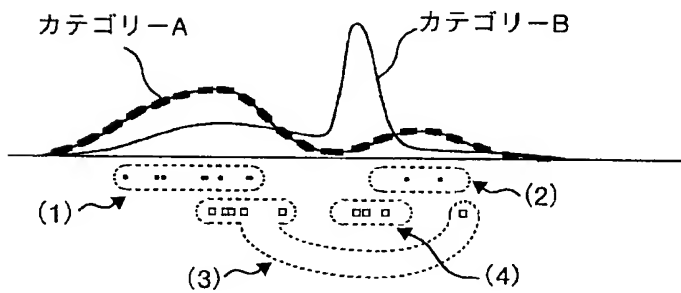


【図4】

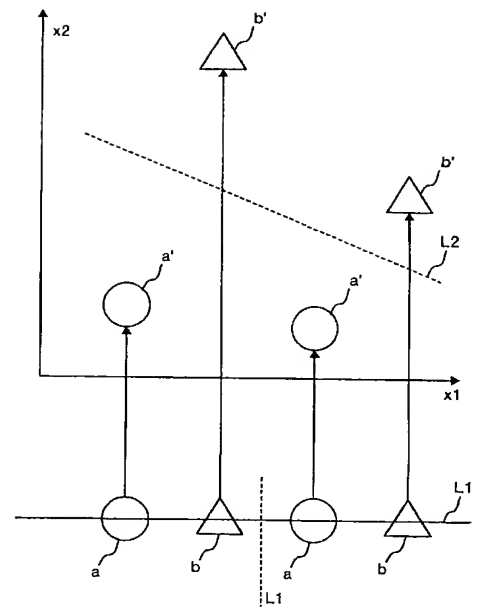
(a)



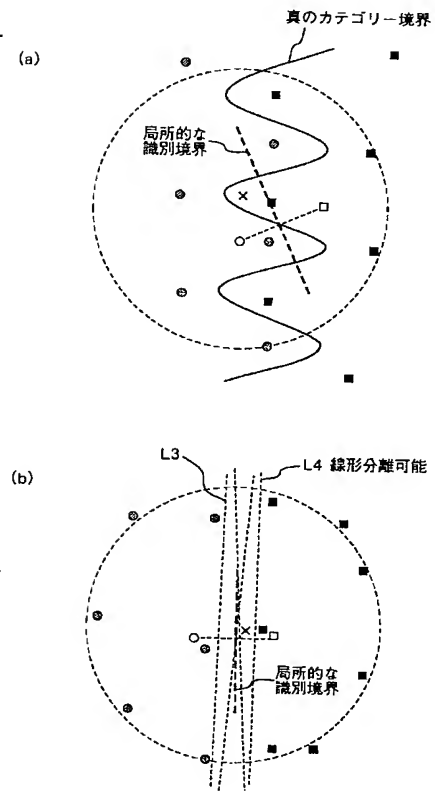
(b)



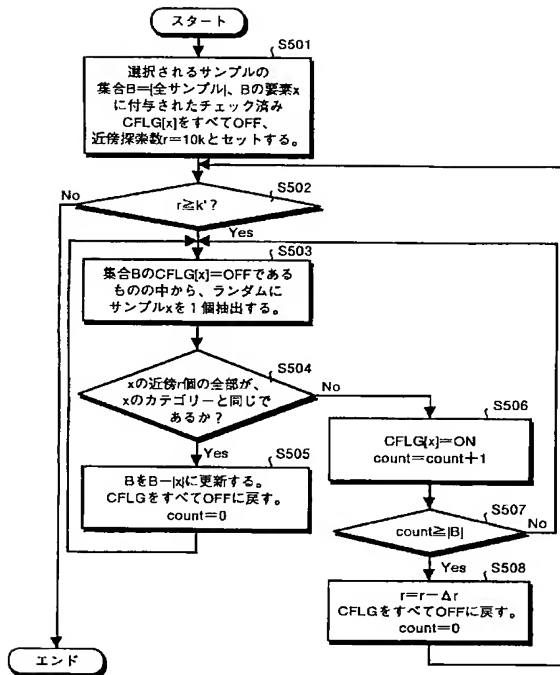
【図10】



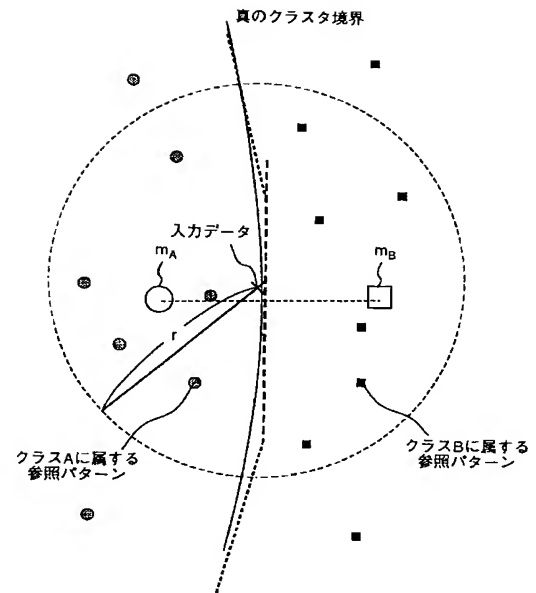
【図11】



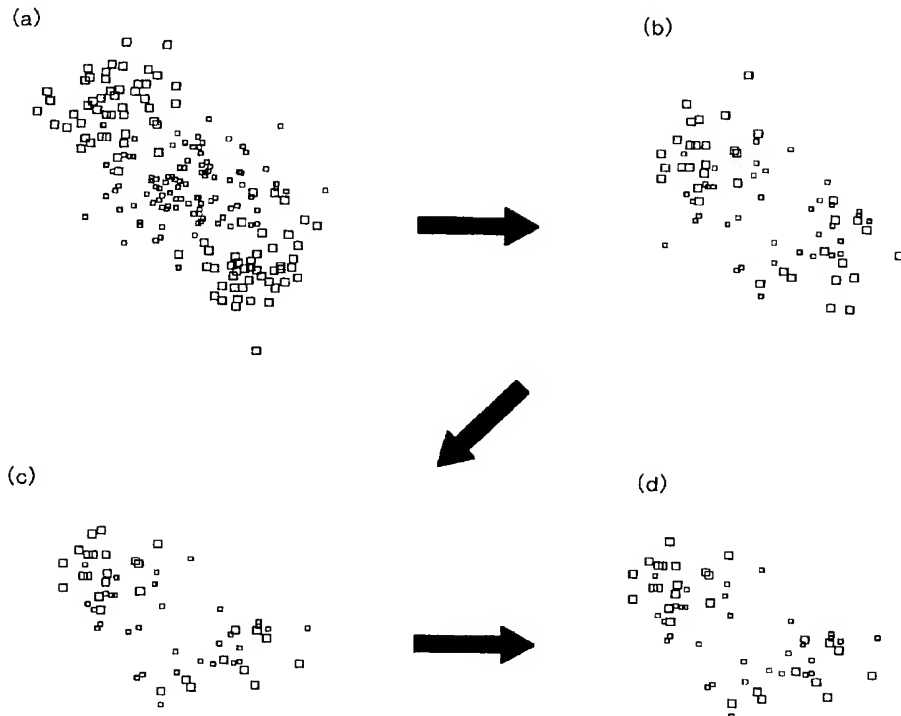
【図5】



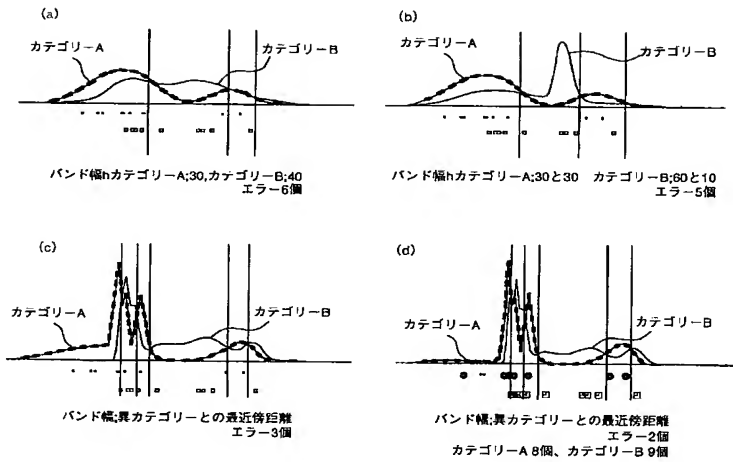
【図9】



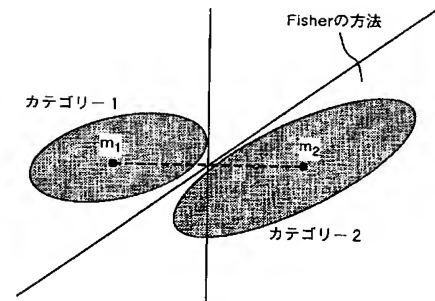
【図6】



【図7】



【図12】



【図13】

